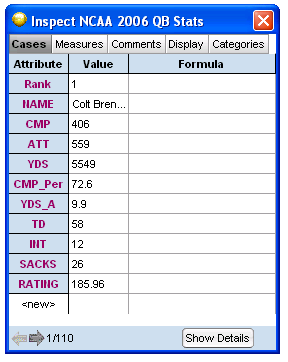
**AP Stat - Introduction to FATHOM!**

**Open Fathom 2**

* Start->Programs->Fathom 2->Fathom 2

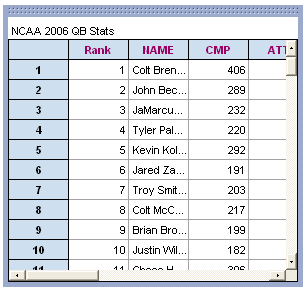
**Opening an Existing File**

* File->Open
* Go to the “STH\_SHARES:” Drive and find the “McNelis” Folder
* Open the SHARE folder, then the STAT folder, and then Open “NCAA 2006 QB Stats”
* In the Blank Fathom Document a collection will drop in.

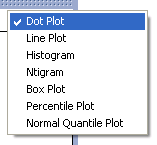


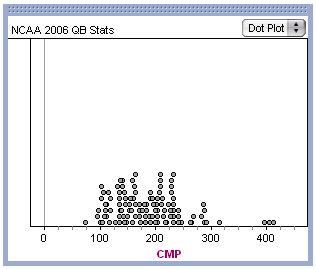
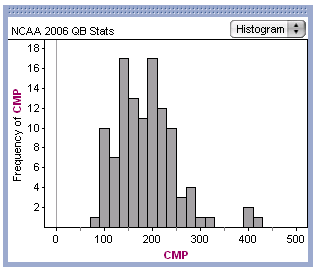
* Double-click on the collection
* The inspect collection window will appear with all the attributes (variables) and the values of a single observation (see picture on right)
* At the bottom LEFT you will see the case number you are in (#1) as well as how many observations are present (110)
* Flip through- Can you identify the 10th quarterback and his # of TD’s?
* Close out of the Inspection window by clicking the red X.

**Opening a spreadsheet**

* Click once on the collection so that it is highlighted
* Fathom uses a lot of drag-drop features
* In the menu, click and hold on the table icon
* Drag the table into the white field
* A spreadsheet should appear with all the attributes on the top of the columns and observations in the cells.
  + If the data does not appear: grab the name of the collection, drag it to the table and drop it in.
* Resize the table so you can see all the columns
* To get rid of a table simple hit delete. The data will not disappear (it is still in the collection. You just closed the table)!
* Bring back the table (recreate it in your document)

**Creating Graphs**

* In the menu click and hold on the graph icon and Drag it into the white field
* A blank graph will appear. An attribute must be dragged into the graph
* From the Table grab the attribute “CMP” and drag it to the bottom axis of the graph
* A dot plot will automatically be drawn
* Change this to a histogram by clicking on the drop-down menu in the top corner and selecting histogram



* To change the bin (bar) width, double click into the white area of the graph. An “Inspect Graph” box will appear.
* Resize this box so you can see all the categories
  + “binAlignmentPosition” tells the program where to start the first bin (where to start your x-axis)
  + “binWidth” tells the program how wide the bins should be
* Change the graph so that the first bin starts at 50 and the width of each bar is 30
* Now the bars go too high- they are off the graph. Fix this by changing the “yUpper” number or **grab the top of the y-axis and drag down** (this is cooler!). Make 25 your upper bound.
* Change the attribute being graphed: grab “ATT” and drag it over “CMP” on the graph. Drop it and it should change the histogram to “ATT”
* Change the start of the first bin to 150, and the bin width to 50. Make sure to adjust the graph so you can see all the bars completely. What is the shape of this distribution?
* Select one bar in the graph. You will notice that the observations that are in that bar are now highlighted in the spreadsheet.

**Copying Graphs to Word Documents**

* To copy a graph to a word document, select the entire graph “ATT”
* In the menu select Edit->**Copy as Picture** or hit **Ctrl+Shift+C**
* Open a word document. In the document select Edit->Paste or hit Ctrl+V
* Delete the picture

**Practice #1**

1. Open a blank word document and put your name at the top of the page.
2. On the next line, put “PRACTICE #1”, then hit enter
3. Now in Fathom, create 2 histograms: one for “YDS\_A” and another for “TD”
4. Copy these into the word document.
5. Save the file as your: lastname.firstname (example: McNelis.Lauren) to your student folder on the computer. Leave the document open.
6. Go back to Fathom and exit out of “NCAA 2006 QB Stats” by clicking the X in the upper right-hand corner of the window in Fathom. (do not exit out of fathom, just the document you were working on!)

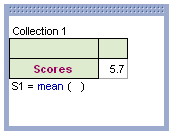
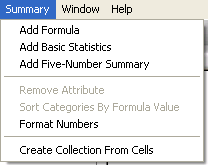
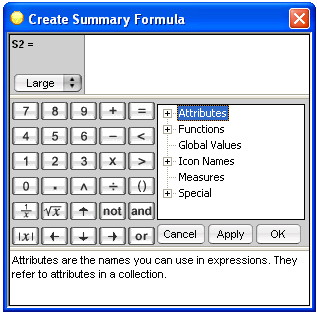
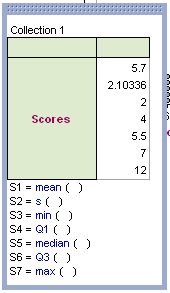
**Practice #2: Creating New Collections**

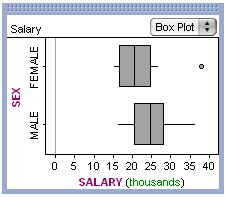
* Open a new, blank fathom document.
* Grab the Collection icon in the menu and drag it to a blank field 
* An empty box will be shown.
* Open a table for the collection (drag the table icon to the blank screen)
* Click on the **<new>** attribute. Rename it “**Scores**”
* Enter the following data:

6, 8, 3, 5, 7, 2, 4, 6, 8, 4, 3, 7, 9, 5, 6, 4, 7, 6, 3, 5, 4, 5, 5, 8, 7, 6, 7, 5, 4, 12

* Create a dot plot of the data. Copy the dotplot to your word document (Start a new page in the document and label with a heading called PRACTICE #2. Put the plot under that heading)
* Save the word document again.

**Creating Summary Tables**

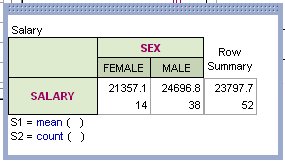
* Grab the Summary icon in the menu and drag it to the blank field 
* A blank summary table will appear. Drag the “Scores” attribute to the left side of the summary table.
* The only statistic given is the mean. To add more statistics select “Summary” in the drop down menu at the top of the page. Select “Add Formula”
* To add the standard deviation, type “s()” into the formula.
* Hit OK, then resize the table so you can see the #’s
* Select the “Summary” menu again and select “Add Five-Number Summary”
* Resize the table

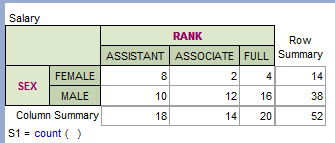


**Importing Data From Another Source**

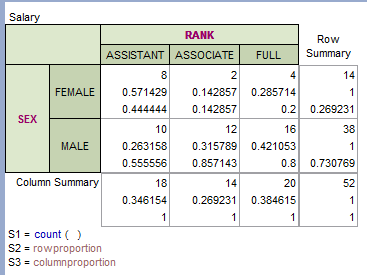
* In a ***new*** Fathom document go to File->Import->Import From File
* In the SHARE folder again, open Salary.TXT
* The collection should appear in the window.

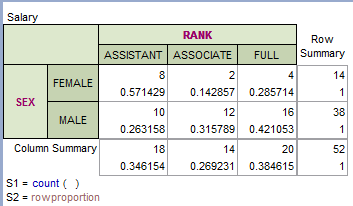
**Comparing Two Separate Groups**

* Open a table and a graph. Make a graph of “Salary” as a boxplot.
* We can add a categorical variable (such as Gender) to the boxplot.
* Grab the attribute “Sex” and drag it to the y-axis of the graph.
* The same can be done with a summary table (see picture at right). Create a summary table with Salary on the left and Gender on the top.

**Creating Two-Way Tables**

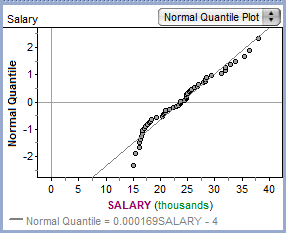
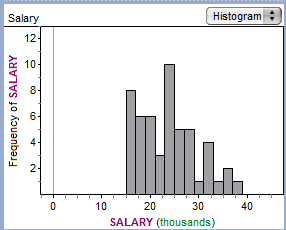
* Pull down a new summary table.
* On the left of the table drag the attribute “SEX”
* On the top of the table drag the attribute “RANK”
* This will create a two-way table since both attributes are categorical.
* To create the conditional distribution of SEX, select the “Summary” menu and select “Add Formula”
* In the formula box type “rowproportion” and select “OK”
* To create the conditional distribution of RANK, select the “Summary” menu and select “Add Formula”
* In the formula box type “columnproportion” and select “OK”



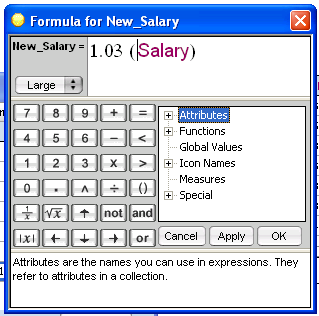


**Practice #3**

1. Make a new page on your word document and label it PRACTICE #3.
2. Create parallel boxplots for SALARY, broken down by DEGREE. Also create summary tables with mean, st dev, and 5# summary for the same situation.
3. Copy and paste both the boxplots and the summary table into your word document.
4. Create a two-way table for “SEX” and “DEGREE” with row and column proportions. Add this to your word doc.
5. Save the word document again!

**Testing for Normality**

* To see if a set of data fits a Normal model use a Normal Quantile Plot.
* Pull down a new graph. Drag in the Attribute “SALARY” into the bottom.
* From the drop-down menu choose “Normal Quantile Plot”
* The more the data fits along the line the closer it will fit a Normal Distribution. For this data a right skew can be seen in the Normal Quantile plot from the points on the left not fitting.



**Adding Formulas**

* Suppose the professors are to get a 3% increasing in pay. We want to see what their new pay will be.
* Double click on the collection “Salary” and the inspect collection box will appear
* In the **<new>** attribute type “New\_Salary”
* Double click in the blank area under “Formula” in the new row. A formula box will appear. Type in “1.03(Salary)” and hit ok.
* The cases should all be entered in automatically.
* Close the fathom document with Salary.

**UNIT 2: Making Scatterplots**

Use the NCAA 2006 collection again (reminder: this can be found in my SHARE folder)

* Put the collection into a blank fathom document. Click once on the collection so it is highlighted.
* Drag down a table so that you can see all the attributes again.

Let’s make a scatterplot of “Attempts” versus “Touchdowns”

* Drag down a new graph. Grab the attribute “ATT” and drag it to the x-axis.
* Grab the attribute “TD” and drag it to the y-axis of the graph.
* You should have a scatterplot of ATT vs. TD



**Finding correlation coefficient**

* Drag down a new summary table.
* Drag the attribute “ATT” to the left row
* Drag the attribute “TD” to the top column
* The correlation coefficient should be stated in the center.

**Finding the LSR line, and the Residual plot**

* While the scatterplot is highlighted (use the same one above with ATT vs. TD), go to the drop-down menu GRAPH and click on “Least-Squares line”
* You will notice that the LSR line has been added to your scatterplot and the equation and r2 are listed down at the bottom of the plot.
* To make the residual plot: Make sure the graph is still highlighted, and go to the menu GRAPH again, and this time click on “Make Residual Plot”
* The residual plot will appear below the scatterplot. Make the entire picture bigger so you can clearly see the residual plot.

**PRACTICE #4**

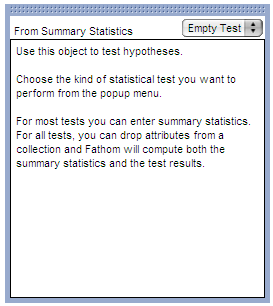
1. Open a new Fathom document.
2. Drag down a new collection.
3. Drag down a table, and label the variables X and Y
4. Enter the values from the table below:

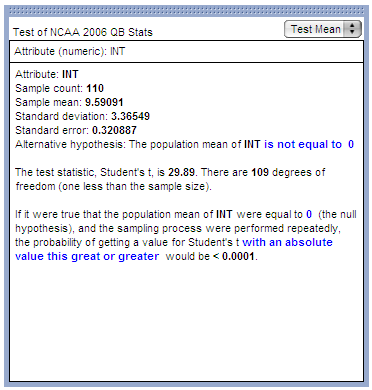
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| x | 1 | 2 | 4 | 5 | 6 | 7 | 2 | 4 | 3 | 9 | 10 | 7 |
| y | 2 | 4 | 3 | 8 | 10 | 9 | 5 | 7 | 9 | 15 | 11 | 10 |

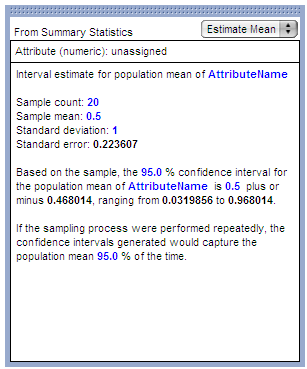
1. Create a scatterplot and a summary table showing the correlation coefficient.
2. Now create the LSR line and find r2. Also, add the residual plot. Make the graph bigger. Add this to your word document (create a new page called PRACTICE #4 and add it under that)
3. Now add the variable of GENDER to your table- just make the first half of the entries female (just type F into the column in the table) and the second half of the entries male (just type M into the table).
4. Make a NEW scatterplot with X vs. Y. Now drag the attribute GENDER into center of that plot. This should give you a scatterplot where there are different symbols for the different genders. Add this to your document.
5. Now create a histogram of the X variable only. Then add GENDER to the Y-Axis of the plot. You should see two histograms, one for each gender. Add this to your word document.
6. Save the document!

**UNIT 6 & 7: Inference on Fathom (confidence intervals and tests)**

* Use the NCAA QB stats collection still
* Have open the table of all the data so that you can see all attributes

**TESTS**

* Drag down a “TEST” from the drag down menu at the top of the page
* In the test, change the drop down menu to “test mean”
* Drag the attribute “INT” and drop it at the top of the test box, where it says “attribute name.”
* You can change your alterative hypothesis by clicking on the blue writing where it says “not equal to.” Change your alternative hypothesis to “less than.”
* You can change your claim by deleting the number “0” and typing in another. Change your claim to 10. Be sure to hit enter after you type it in.
* You should now see your test statistic of -1.275, a df of 109, and at the bottom a p-value of 0.10 (so we would fail to reject).
* To test two means, you would select “compare means” from the top drop down menu, and then put in both attributes. Test whether the average TDs is less than the average SACKS.
* Confirm that you get a test statistic of -4.084, a df of 217.976, and a p-value < 0.0001.
* You can do the same thing for proportions, however you have to have a categorical variable, with only 2 outcomes.

**INTERVALS**

* Drag down a “ESTIMATE” from the drag down menu at the top of the page
* In the estimate box, change the drop down menu to “estimate mean”
* Now drag the attribute “YDS\_A” into the attribute name at the top of the box.
* You will see that it automatically does 95% confidence interval (in blue in the center of the box). You can change this by typing in a new confidence level. Change this to 99% confidence. Be sure to hit enter.
* You will see the confidence interval at the end of the middle paragraph. It should say (**6.87409, 7.38227**).
* To make a confidence interval with 2 means, you would select difference of means. Find the interval for the difference between the averages of TDs and SACKS.
* Confirm that you get a 95% interval of (-7.22443, -2.52102).
* You can do proportions too. Just be sure that you have a categorical variable where there are only 2 outcomes.

**PRACTICE #5**

Open a new collection and a new table. Add the data for the following 2 variables:

RESPONSE: y, y, y, n, n, y, y, y, y, n, n, n, y, y, y, y, y, n, n, n, n, y, y, y, y, y, y

AGE: 12, 15, 16, 14, 15, 16, 17, 14, 15, 16, 14, 17, 18, 16, 15, 14, 11, 17, 12, 14, 11, 12, 14, 15, 16, 16, 17

1. Test whether the percent of Yes responses is greater than 60%. Copy your test box into your word document under a new heading of PRACTICE #6.
2. Find a 97% confidence interval for the average age of the respondents. Copy your interval box into your word document.
3. Save the word document again. Close it. Drop it into my drop folder. Close Fathom.